Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Journal of Molecular Liquids 314 (2020) 113612

Contents lists available at ScienceDirect

Journal of Molecular Liquids

journal homepage: www.elsevier.com/locate/molliq

Immunoinformatics characterization of SARS-CoV-2 spike glycoprotein for prioritization of epitope based multivalent peptide vaccine

Saba Ismail [1], Sajjad Ahmad [1], Syed Sikander Azam *

Computational Biology Lab, National Center for Bioinformatics (NCB), Quaid-i-Azam University, Islamabad 45320, Pakistan

ARTICLE INFO

Article history:
Received 5 April 2020
Received in revised form 14 May 2020
Accepted 14 June 2020
Available online 17 June 2020

Keywords:
COVID-19
SARS-CoV-2
Spike glycoprotein
Vaccine
Immuno-informatics
Multi-epitope peptide vaccine construct
Molecular dynamics simulation

ABSTRACT

The COVID-19 pandemic caused by SARS-CoV-2 is a public health emergency of international concern and thus calling for the development of effective and safe therapeutics and prophylactics particularly a vaccine to protect against the infection. SARS-CoV-2 spike glycoprotein is an attractive candidate for a vaccine, antibodies, and inhibitors development because of the many roles it plays in attachment, fusion and entry into the host cell. In the present investigation, we characterized the SARS-CoV-2 spike glycoprotein by immunoinformatics techniques to put forward potential B and T cell epitopes, followed by the use of epitopes in construction of a multi-epitope peptide vaccine construct (MEPVC). The MEPVC revealed robust host immune system simulation with high production of immunoglobulins, cytokines and interleukins. Stable conformation of the MEPVC with a representative innate immune TLR3 receptor was observed involving strong hydrophobic and hydrophilic chemical interactions, along with enhanced contribution from salt-bridges towards inter-molecular stability. Molecular dynamics simulation in aqueous milieu aided further in interpreting strong affinity of the MEPVC for TLR3. This stability is the attribute of several vital residues from both TLR3 and MEPVC as shown by radial distribution function (RDF) and a novel axial frequency distribution (AFD) analytical tool. Comprehensive binding free energies estimation was provided at the end that concluded major domination by electrostatic and minor from van der Waals. Summing all, the designed MEPVC has tremendous potential of providing protective immunity against COVID-19 and thus could be considered in experimental studies.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

In December 2019, a new strain of coronavirus emerged in Wuhan city of Hubei province in China and has since disseminated globally. The virus belongs to clade B of family Coronaviridae in the order Nidovirales, and genera Betacoronavirus and caused pulmonary disease outbreak [1,2]. It is positive-sense RNA, enveloped and non-segmented virus and named as SARS-CoV-2 as it shares 82% genome homology with SARS coronavirus (SARS-CoV) [3,4]. SARS-CoV-2 causes coronavirus disease-19 (COVID-19) and evidence suggest a zoonotic origin of this disease [5]. Though the zoonotic transmission is not completely understood but facts provide the ground that it proliferates from the seafood market Huanan in Wuhan and human-to-human transmission resultant into the exponential increase in number of cases [6,7]. As of May 12, 4,320,202 cases are reported worldwide with 291,545 deaths and 1,570,090 recovered. Among the active cases, 2,458,567 are currently infected, 2,412,235 (98%) are in mild conditions and 46,332 (2%) are seriously ill. Among the 1,861,635 closed cases, 1,570,090 (84%) are recovered whereas 291,545 (16%) die. On March 11, the World Health Organization (WHO) affirmed COVID-19 as a pandemic (https://www.worldometers.info/coronavirus/).

SARS-CoV-2 utilizes a highly glycosylated, homotrimeric class I viral fusion spike protein to enter into host cells [8]. This protein is found in a metastable pre-fusion state which goes through structural readjustments facilitating fusion of the viral membrane to the host cell [9–11]. The binding of S1 subunit to the host angiotensin converting enzyme (ACE) initiates the fusion process and during this event the pre-fusion trimeric structure disrupts resulting in S1 subunit dispersion and stabilization of the S2 subunit to achieve a post-fusion conformation [12]. The receptor-binding domain (RBD) of S1 unit adopts a hinge-like conformation to temporarily hide or expose receptor binding residues for interaction with the host cell receptor [11]. Down and up conformation states are recognized where former is related to the receptor inaccessible state and the later one explains receptor accessible state and considered as less stable [13–16]. This critical role of the spike protein makes it an important candidate for neutralization by antibodies, and detailed atomic level studies of the pre-fusion spike structure are important in the design and development of a vaccine [17–21]. Current data indicates that both SARS-CoV-2 and SARS-CoV

* Corresponding author at: National Center for Bioinformatics, Quaid-i-Azam University, Islamabad 45320, Pakistan.
E-mail address: ssazam@qau.edu.pk (S.S. Azam).
[1]Contributed equally to this work as the first author.

https://doi.org/10.1016/j.molliq.2020.113612
0167-7322/© 2020 Elsevier B.V. All rights reserved.

spike share the same ACE2 as a binding receptor [22,23]. Interestingly, ACE2 binds to SARS-CoV-2 spike with ~15 nM affinity, about 10–20 folds higher than ACE2-SARS-CoV spike binding [24]. One possible reason for SARS-CoV-2 human-to-human transmission is SARS-CoV-2 spike's high affinity for human ACE2 [25]. Series of cellular immune and humoral responses can be triggered by SARS-CoV-2 infection [26]. Immunoglobulin G (IgG) and IgM can be noticed after 2 weeks of infection onset which are specific antibodies to SARS-CoV-2. High titers of neutralizing antibodies and SARS-CoV-2 specific cytotoxic T lymphocyte responses have been identified in the patients who cleared the SARS-CoV-2. This phenomenon clearly suggests that both cellular and humoral immune reactions are vital in clearing the SARS-CoV-2 infection [26–30].

The study presented, herein, is an attempt to get insights about antigenic determinants of SARS-CoV-2 spike glycoprotein and highlight all antigenic epitopes [31] of the spike that can be used specifically for the design of a multi-epitope peptide vaccine construct (MEPVC) [32] to counter COVID-19 infections. The epitopes predicted by immunoinformatics techniques were fused together as well as to β-defensin adjuvant [33,34] to boost the antibody production and long-lasting immunological responses. Blind docking protocol was implemented to gain stable MEPVC binding conformation with respect to a representative TLR3, an important innate immune receptor in viral immunity [35]. The docked solution was further utilized in dynamics simulations to study the structural dynamics and stability of the complex [36]. Finally, to confirm intermolecular interactions and point all noteworthy residues crucial for intermolecular stability, binding free energies [37] for the system were estimated. In a nutshell, this study reflects excellent outcomes for experimental vaccinologists to develop a vaccine for prevention from pandemic COVID-19 infections.

2. Materials and methods

Complete flow of the *in silico* designed methodology for MEPVC targeting SARS-CoV-2 spike is demonstrated in Fig. 1.

2.1. Epitopes mapping for spike protein

The spike glycoprotein amino acid sequence was retrieved from NCBI SARS-CoV-2 data hub and considered first in the epitope mapping phase, where T cell epitopes derived from B cell were predicted using immune epitopes database (IEDB) [38]. Linear B cell epitopes were mapped using Bepipred Linear Epitope Prediction 2.0 [39] and those with score >0.5 were subjected to T cell epitopes identification step. The epitopes were projected for association with reference set of major histocompatibility complex (MHC): MHC class I [40] and MHC class II [41] alleles sorted on percentile score basis. Epitopes with lowest percentile score are strong binders and were considered only. The selected epitopes were then used in MHCPred 2.0 [42] to decipher their binding affinity potential for predominant HLA II DRB*0101 and only those with IC_{50} value <100 nM were categorized as excellent DRB*0101 binders [43]. VirulentPred [44] was employed next to reveal virulent nature of the epitopes setting the cut-off to 0.5. Antigenic epitopes were highlighted by VaxiJen 2.0 [45]. Allergenic epitopes were discarded through AllerTop 2.0 [46] and toxic potential of non-allergic epitopes was evaluated *via* ToxinPred [47]. The non-toxic epitopes were lastly investigated for their ability to induce IFN-γ using an IFN epitope server [48]. Conservation across the world population of the final set of epitopes was done through IEDB epitope conservation analysis tool [49].

2.2. MEPVC designing and post analysis

All filtered epitopes were linked together through AAY linkers [50] to design a multi-epitope peptide (MEP). The resultant peptide was further linked to an immunological β-defensin (an adjuvant) to construct a MEPVC and in this way, immunogenicity can be enhanced. The physicochemical properties of designed MEPVC were predicted by ProtParam tool [51] of EXPASSY server. The three dimensional (3D) structure of the MEPVC was modeled *ab initio* by 3Dpro of SCRATCH protein server [52]. Following, loop modeling was done in the 3D structure of MEPVC *via* GlaxyLoop [53] from GlaxyWeb and subsequently refined through GalaxyRefine [54]. Disulfide engineering was applied to the MEPVC refined model *via* Design 2.0 [55] as disulfide bonds strengthen structure stability. The MEPVC sequence was translated reversibly for optimization of codon usage according to *Escherichia coli* K12 expression system in order to get high expression rate [56]. For this, Java Codon Adaptation Tool (JCat) [57] was employed and expression rate of the cloned MEPVC was measured by codon adaptation index (CAI) value. SnapGene (https://www.snapgene.com/) was used to clone the optimized MEPVC cDNA into pET-28a (+) expression vector.

2.3. In silico immune profiling of MEPVC

Immunogenic potential of the MEPVC was conducted *in silico* using the C-ImmSim server [58,59]. The server used machine learning techniques along with position-specific scoring matrix (PSSM) for estimate of the human host immune system response to the antigen. The immune system responds from three sites: bone marrow, lymph nodes and thymus. The input parameters for the immune simulations are as follows: number of steps (100), volume (10), random seed (12345), HLA (A0101, A0101, B0702, B0702, DRB1_0101, DRB1_0101), number of injection set to 1. All remaining parameters were treated as default.

2.4. Molecular docking of MEPVC

The MEPVC affinity for an appropriate immune receptor as an agonist was checked in the step of molecular docking [60]. TLR3 available under PDB id of 1ZIW was retrieved and used as a receptor molecule. TLR3 also named CD283 is a pattern recognition receptors (PPRs) protein and is a transmembrane [61]. It detects viral infection associated dsRNA and evokes the activation of interferon regulatory transcription factor (IRF3) and Nuclear Factor kappa light-chain enhancer (NF-kB) [62]. Unlike other TLRs, TLR3 uses TIR-domain containing adapter inducing interferon-β (TRIF) as a primary adapter [63]. IRF3 eventually induces the development of type I interferons leading to the activation of both innate and adaptive immune systems [64]. The TLR3 receptor and MEPVC were used in a blind docking approach through an online PATCHDOCK server interface [65]. The interacting molecules were docked according to the shape complementarity principle. The clustering RMSD is allowed to default 4.0 Å. The output docked solutions were refined for interactions with Fast Interaction Refinement in Molecular Docking (FireDock) server [66]. The refined complexes were examined and one with lowest global energy was considered as top ranked. The opted complex was subjected to in-depth MEPVC conformation visualization with respect to the TLR3 using UCSF Chimera 1.13.1 [67].

2.5. Molecular dynamics (MD) simulation

MD simulation was applied to the selected top complex for 100-ns to understand complex dynamics and stability for practical applications. This assay was categorized into three phases: (i) parameters file preparation (ii) pre-processing, and (iii) simulation production [68]. In first phase using an antechamber module of AMBER16 [69], complex libraries and set of parameters for TLR3 and MEPVC were generated. The complex system was solvated into 12 Å TIP3P solvation box achieved through Leap module of AMBER. The intermolecular and intramolecular interactions of the system were determined by ff14SB force field [70]. Counter ions in the form of Na^+ were added to the

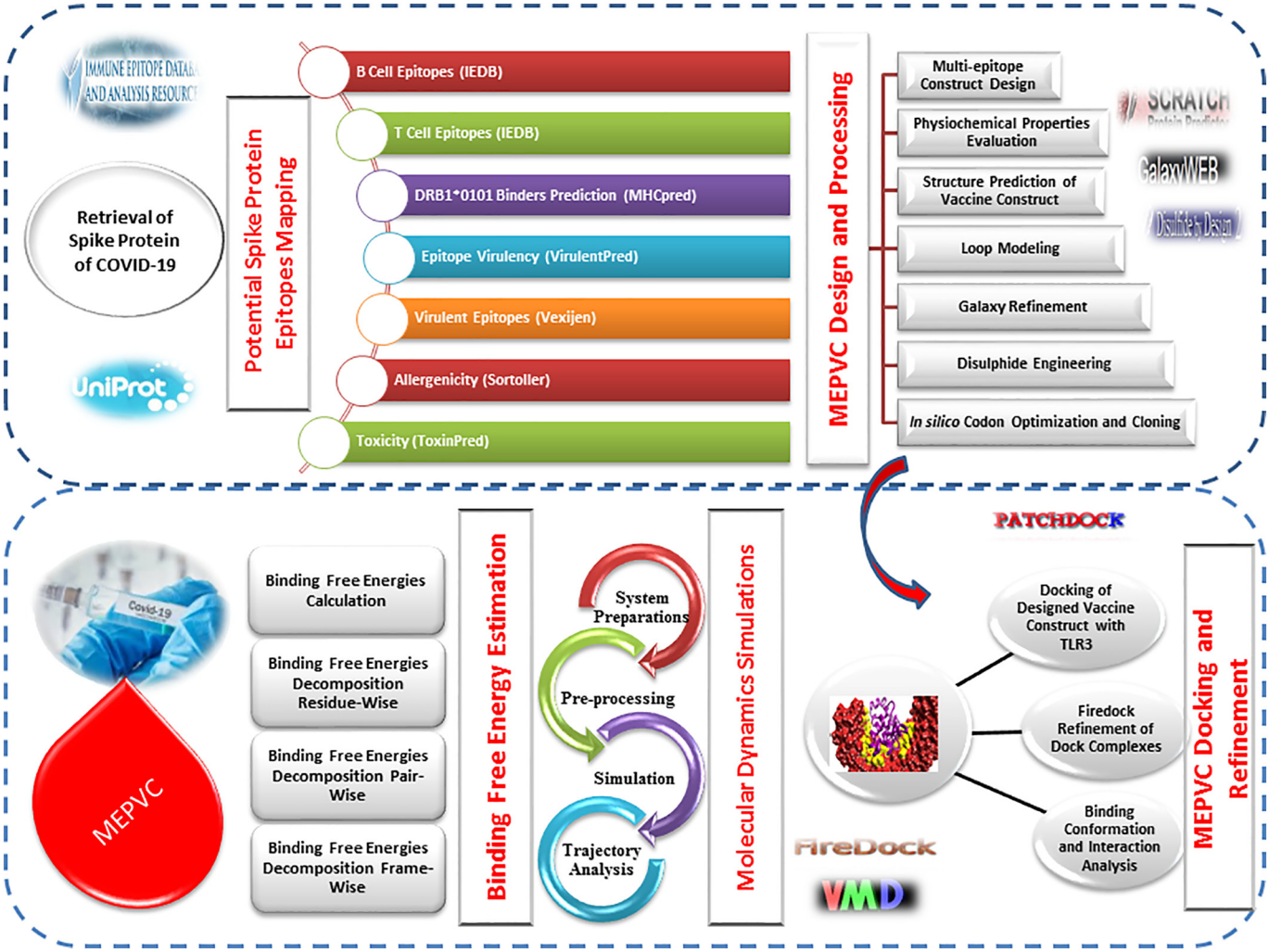

Fig. 1. Computational approach adopted for the design of a SARS-CoV-2 spike protein based MEPVC.

system for charge neutralization. In the system pre-processing stage, complex energy was optimized through several rounds: minimization of complete set of hydrogen atoms for 500 steps, solvation box energy minimization for 1000 steps with 200 kcal/mol – Å^2 restraint on the remaining system, minimization of complete set of system atoms again for 1000 steps with 5 kcal/mol –Å^2 applied restraint on carbon alpha atoms, and 300 minimization steps on non-heavy atoms with restraint of 100 kcal/mol –Å^2 on other system components. The complex system then underwent a heating step where the complex was heated gradually to 300 K through NVT ensemble, maintained through Langevin dynamics [71] and SHAKE algorithm [72] for restraining hydrogen bonds. Complex equilibration was achieved for 100-ps. System pressure was maintained using NPT ensemble allowing 5 kcal/mol – Å^2 energy restraint on Cα atoms. In the simulation production, trajectories of 100-ns were produced on time scale of 2-fs. Non-bounded interactions were differentiated by describing cut-off distance of 8.0 Å.CPPTRAJ module [73] was lastly used for statistical computation of different structure parameters to probe complex stability. The MD simulation frames were analyzed and visualized in Visual Molecular Dynamics (VMD) 1.9.3 [74].

2.6. Axial frequency distribution analysis

Axial frequency distribution or simply AFD [75] is a novel analytical technique run on simulation trajectories to access ligand 3D conformation with respect to a reference receptor atom. Such local structural movements are not captured by any other available technique. AFD is mathematically presented by Eq. (I),

$$AFD = \sum_{i=1,j=1}^{k,l} mi,j \tag{I}$$

where, i and j are ligand atom coordinates on the X and Y axis with cut-off value k and l, respectively. The mi,j sums interactions frequency that falls in the coordinate (i,j).

2.7. Calculating binding free energy and binding entropy

The interaction energy and solvation free energy for TLR3 receptor, MEPVC, TLR3-MEPVC complex were calculated utilizing the MMPBSA.py module [76] of AMBER16. Also, an average of the above was estimated as system net binding free energy. The binding free energy was computed through the MM-PBSA method and its counterpart MM-GBSA with objective to derive the energy difference between bound and unbound states of the same system [77]. Mathematically, the free energy can be calculated through Eq. (II),

$$\Delta G_{bind,solv} = \Delta G_{bind,vaccum} + \Delta G_{solv,TLR3-MEPVC} - (\Delta G_{solv,MEPVC} + \Delta G_{solv,TLR3} \tag{II}$$

For all three states of the system, the solvation energy was estimated by solving either Generalized Born (GB) or Poisson Boltzman (PB)

equation and thus it will give electrostatic energy. It also allows the addition of empirical terms for hydrophobic contributions as shown in Eq. (III).

$$\Delta G_{solv} = G_{electrostatic,\epsilon=80} - G_{electrostatic,\epsilon=1} + \Delta G_{hydrophobic} \quad \text{(III)}$$

Calculation of the average interaction energy between TLR3 and MEPVC gives to delta- ΔG_{vacuum} (Eq. (IV)).

$$\Delta G_{vacuum} = \Delta E_{moleular\ mechanics} - T.\Delta S \quad \text{(IV)}$$

The binding entropy was calculated using the normal mode analysis method implemented in AMBER16 which is based on Quasi-Harmonic approximation and run through PTRAJ program in AmberTools. As such calculations are computationally expensive, only 10 frames of the simulation trajectories were considered.

3. Results and discussion

3.1. Prioritization of potential epitopes

The SARS-CoV-2 spike protein was targeted for MEPVC designing because of the many filters it fulfilled and required for a potential vaccine candidate. First, it does not share any significant homology to the human host, and as such chances of autoimmune responses are negligible [78]. Second, the protein is also not found to have any sequence identity to the mouse proteome and thus accurate immunological findings can be deciphered from *in vivo* mice experimentations [79]. This spike protein only harbored one transmembrane helix ensuring the wet lab protein cloning and expression for antigen analysis easy [80]. Antigenicity is another factor that makes this candidate highly suitable for vaccine designing as this allows efficient binding to the products of host immune system [81]. Further, this protein is strongly adhesive which makes it an excellent target for creation of adhesion based vaccine [82]. Lastly, all the sequences of SARS-CoV-2 spike protein are highly conserved thus a vaccine based on its sequence will be highly likely to have broad spectrum immunological implications [83]. Prioritization of potential epitopes for the SARS-CoV-2 spike protein commenced with the mapping of B cell epitopes that predicted total of 34 epitopes of variable length ranging from 1 to 62 (S-Table 1). The average score predicted for these B cell epitopes is 0.470, maximum (max) of 0.696 and minimum (min) of 0.188 (S-Fig. 1). Each B cell epitope was then analyzed in MHC-1 alleles binding regions prediction [84]. A set of reference alleles to which these epitopes interact with are: HLA-A*01:01, HLA-A*02:01, HLA-A*02:01, HLA-A*02:03, HLA-A*02:03, HLA-A*02:06, HLA-A*02:06, HLA-A*03:01, HLA-A*03:01, HLA-A*11:01, HLA-A*11:01, HLA-A*23:01, HLA-B*08:01, HLA-A*23:01, HLA-A*24:02, HLA-A*24:02, HLA-A*26:01, HLA-A*26:01, HLA-A*30:01, HLA-A*30:01, HLA-B*57:01, HLA-A*30:02, HLA-A*31:01, HLA-B*58:01, HLA-A*32:01, HLA-A*33:01, HLA-A*33:01, HLA-A*68:01, HLA-A*68:01, HLA-A*68:02, HLA-A*68:02, HLA-A*30:02, HLA-B*07:02, HLA-B*51:01, HLA-B*07:02, HLA-B*08:01, HLA-B*15:01, HLA-B*15:01, HLA-B*35:01, HLA-A*31:01, HLA-B*35:01, HLA-B*40:01, HLA-B*40:01, HLA-B*44:02, HLA-B*44:02, HLA-B*44:03, HLA-B*44:03, HLA-B*51:01, HLA-A*01:01, HLA-B*53:01, HLA-B*53:01, HLA-B*57:01, HLA-A*32:01, and HLA-B*58:01. These alleles have >97% population coverage. The predicted epitopes were then screened and stringent criteria of lowest percentile score was used to choose the excellent binders. Afterward, the B cell epitopes were simultaneously run in MHC-II alleles binding [85]. Likewise MHC-I, reference set of MHC-II binding were: HLA-DRB4*01:01, HLA-DRB1*04:01, HLA-DRB1*04:05, HLA-DRB1*07:01, HLA-DRB1*09:01, HLA-DRB1*11:01, HLA-DRB1*03:01, HLA-DRB1*13:02, HLA-DRB1*15:01, HLA-DRB3*01:01, HLA-DRB1*12:01, HLA-DRB3*02:02, HLA-DRB1*08:02, HLA-DRB1*01:01, and HLA-DRB5*01:01. The MHC-II predicted epitopes were also filtered on basis of percentile score and then cross checked with the selected MHC-I allele and those common in both classes were considered only which were 50 in numbers. The shortlisted common MHC-I and MHC-II epitopes then subjected to antigenicity check. In this check, ability of the filtered B cell derived T cell epitope's to evoke and bind to products of adaptive immunity was evaluated. This yielded 38 epitopes all of which have strong ability to bind to the most prevalent DRB*0101 with average IC_{50} score of 35.6552, max of 98 and min of 0.89.The antigenic epitopes then underwent allergenicity check to discard allergic peptides that may cause allergic reactions [86]. This resulted into 31 epitopes. Non-toxic epitopes were 7 whereas 6 were IFN-gamma producer (Fig. 2). The set of epitopes obtained at different stages of epitope mapping phase is tabulated in S-Table 2. These epitopes appear to provide coverage to 98% of the world population (Fig. 3). The coverage can split as: 98.54% in East Asia, 96.06% in Northeast Asia, 97.46% in South Asia, 95.83% in Southeast Asia, 94.29% in Southwest Asia, 98.53% in Europe, 92.81% in East Africa, 89.82% in Central Africa, 96.45% in West Africa, 97.86% in North Africa, 94.81% in South Africa, 99.09% in West Indies, 98.74% in North America, 25.01% in Northern America, 90.91% in South America, and 95.58% in Oceania.

3.2. MEPVC designing

Prioritized T cell epitopes derived from B cells were fused together tandemly by AAY linkers to make a MEP. AAY linker avoids formation of junctional epitopes and as such enhance epitope presentation [87]. To the N-terminal of MEP, EAAAK linker was added to attach β-defensin as an adjuvant leading to the design of a MEPVC. The MEPVC is schematically shown in Fig. 4A. MEPVC offers many advantages compared to a separate antigenic peptide. Such vaccines induce both CD4+ and CD8+ responses and the antigens optimization is optimal. EAAAK is a rigid spacer and allows separation of the attached domain and promoting efficient immune processing of the epitopes [88]. β-defensins are potent immune adjuvants as they are capable of significantly enhancing production of lymphokines resulting in antigen-specific Ig production and T cell-dependent cellular immunity. The sequence of MEPVC is GIINTLCKYYCRVRGGRCCVCSCCPKEEQIGKCS TRGRKCCRRKKECAAKYAWNRKCISACYIAPGQTGKICCYPRRARSVCSACYT VYDPCQPCAAYVYDPLCPELCAYCKNHTSCDV.

3.3. MEPVC has appropriate biophysicochemical properties

Physicochemical properties of the MEPVC were evaluated in order to assist experimentalists in the field to set up *in vitro* and *in vivo* experiments accordingly. The length of MEPVC is spanned across 110 amino acids and has molecular weight of 13.30 kDa. Vaccine construct with weight <110 kDa is generally believed to effective vaccine target because of its easy purification. Theoretical pI of MEPVC is 9.8 and aids in locating MEPVC on 2D gel. MEPVC aliphatic index is 69.08 projecting the vaccine thermostable at different temperatures. The total number of negatively charged and positively charges residues are 8 and 22, respectively. The grand average of hydropathicity (GRAVY) score computed for the MEPVC is −0.545, illustrating hydrophilic nature of the protein and is likely to interact with water molecules. The estimated half-life of MEPVC in mammalian reticulocytes, *in vitro* is 30 h, yeast, *in vivo* is >20 h, and *E. coli*, *in vivo* higher than 10 h. The antigenicity of the MEPVC was cross-checked and predicted highly antigenic with value of 0.69. Total entropy of the protein is 17.0170 which is considered ideal and also the vaccine has no transmembrane helices (alpha helical transmembrane protein, 0.0474783 and beta barrel transmembrane protein, 0.0060384) hence no difficulties can be anticipated in cloning and expression analysis. The predicted solubility upon overexpression of MEPVC is 0.965751 reflecting higher solubility of MEPVC.

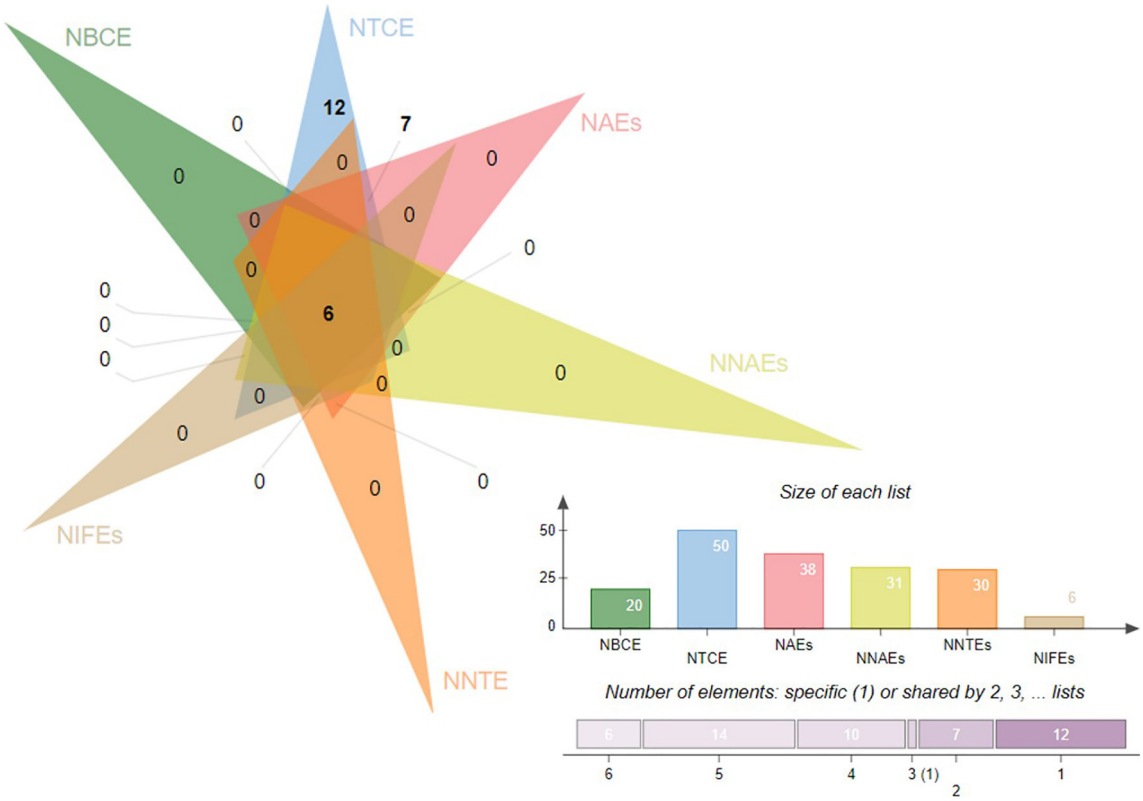

Fig. 2. Venn diagram for the number of different categories of epitopes predicted at the epitope mapping stage. NBCE (number of B cell epitopes), NTCE (number of T cell epitopes), NAEs (number of antigenic epitopes), NNAES (number of non-allergic epitopes), NNTEs (number of non-toxic epitopes), and NIFEs (number of IFN-gamma epitopes).

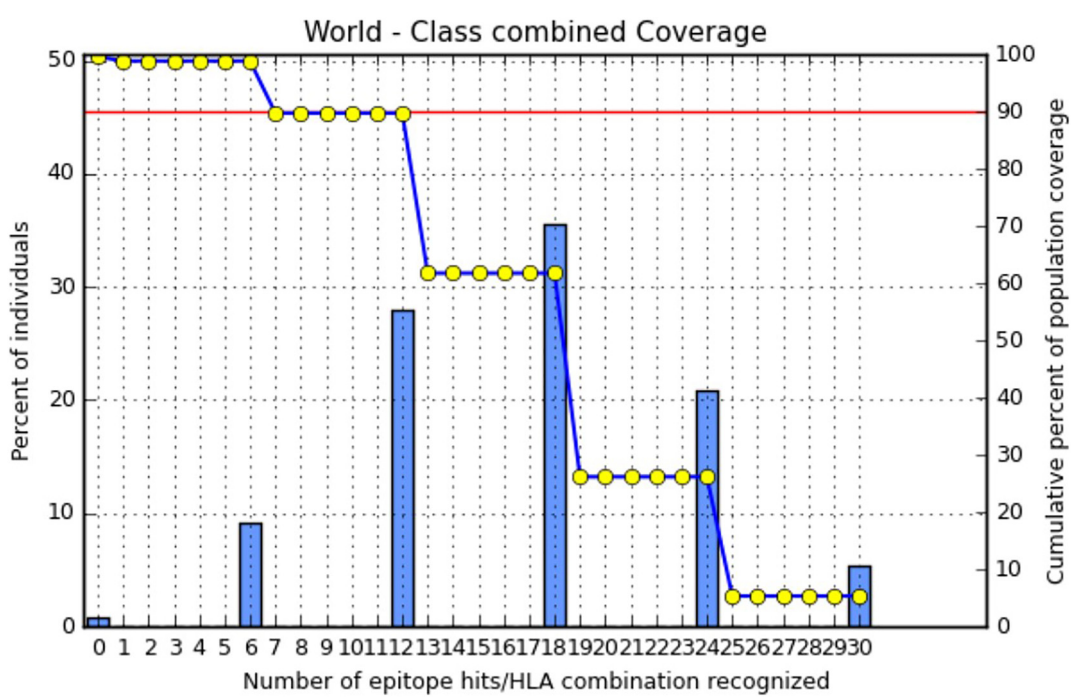

Fig. 3. World population coverage of the final set of predicted epitopes.

3.4. Modeling and refinement of MEPVC

The 3D model of the MEPVC was constructed using *ab initio* 3Dpro predictor as no appropriate template was available for homology modeling and threading methods. The 3D structure of MEPVC is shown in Fig. 4B. The structure secured 85.4% of residues in the Ramachandran favored, 12.6% in additionally allowed, 1.9% in generously allowed, and 0% in disallowed region. As the predicted

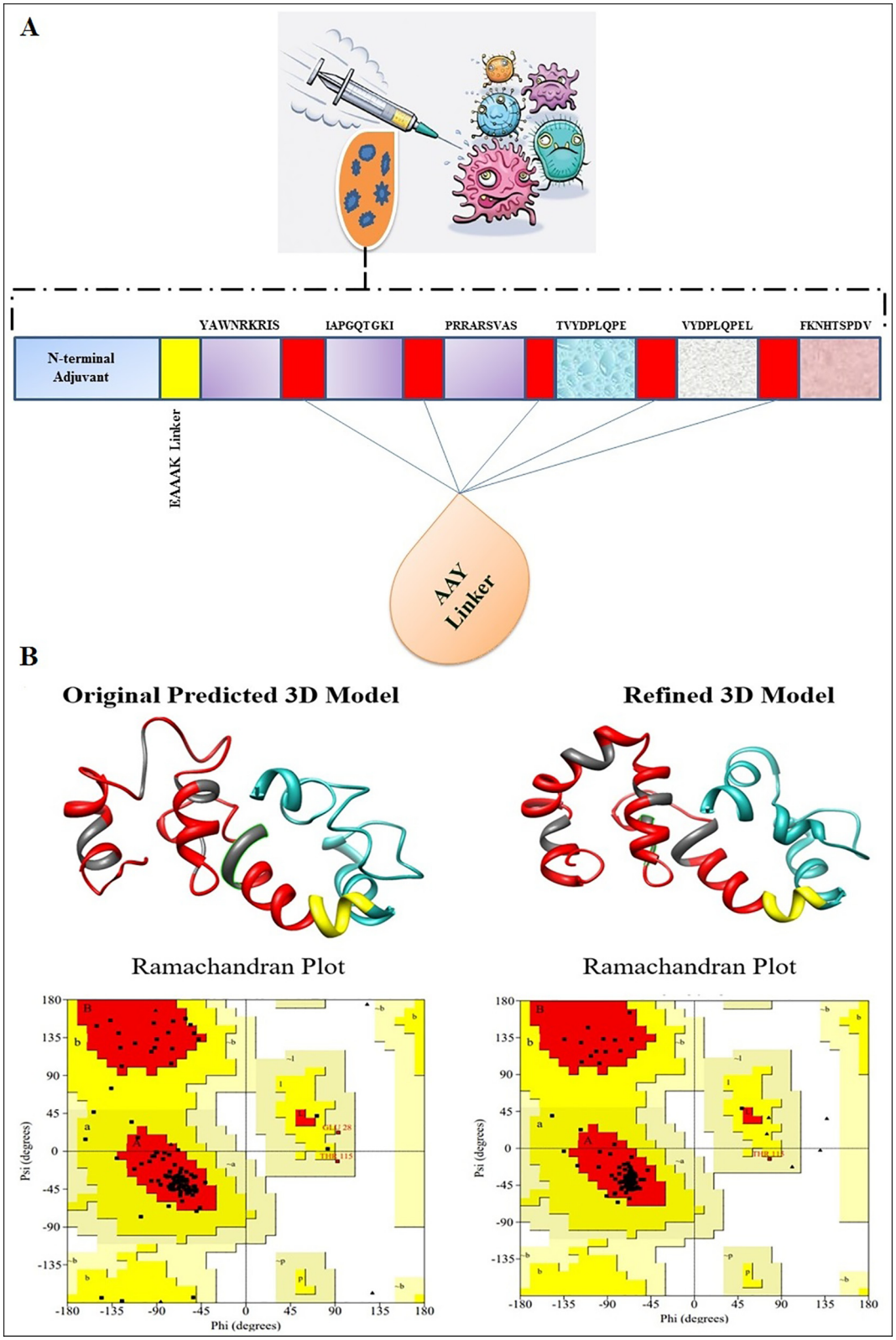

Fig. 4. A. Schematic depiction of the MEPVC. B. The original predicted 3D MEPVC structure and refined along with respective Ramachandran plots. AAY linkers are shown in red while epitopes are in coal and yellow is for EAAAK linker. Cyan color represents the β-defensin adjuvant. In the Ramachandran plot, the torsion angles are shown by black squares dispersed across the core secondary structures (colored as red). The allowed regions can be understood by color yellow, generously allowed by pale yellow and disallowed by white region. The top right, top left, bottom right and bottom left represent quadrants for left handed alpha helices, beta sheets, right handed alpha helix, and no elements, respectively.

MEPVC unit has number of loop regions that need to be modeled properly before moving forward. In total, five sets of residues: Alas7-Lys32, Ile63-Gly69, Cys73-Arg77, Thr87-Pro102, and Asn113-Val119 were loop modeled. The loop modeled structure increased the Ramachandran favored residues percentage to 92.3%, residues in allowed region reduced to 6.7%, residues of generously allowed region

to 1.0% and disallowed remained to 0%. The structure was subjected to structure perturbations and relaxations to obtain a refined model. Among the generated structures (S-Table 3), the first model was selected as it has improved Rama favored score, lowest stable galaxy energy of 0.96, improved clash score of 23.1 and good MolProbity value. Similarly, the structure lacks poor rotamers in contrast to the original structure. The Ramachandran statistics for the refined 3D structure are in following order: Ramachandran favored residues (93.2%), additionally allowed region (5.8%), generously allowed region (1.0%) and disallowed region (0%). The *Z*-score of the refined MEPVC is −4.3 and within the score range of the same size protein in structure databases (S-Fig. 2).

3.5. Disulfide engineering and in silico cloning of the MEPVC

Further, disulfide engineering of the MEPVC was performed in order to optimize molecular interactions and confer considerable stability by attaining precise geometric conformation [89,90]. Eight pairs of residues were selected to be replaced with cysteine amino acids. These pairs are: Gln7-Ala19 (χ3 angle,+118, energy value, 4.20 kcal/mol), Cys18-Leu21 (χ3 angle,+84.35, energy value, 3.69 kcal/mol), Lys44-Ala47 (χ3 angle, +74.17, energy value,5.59 kcal/mol), Arg57-Ala61 (χ3 angle,+122.71, energy value,6.14 kcal/mol), Ala72-Ala85 (χ3 angle,-62.92, energy value,4.40 kcal/mol), Ala73-Ala82 (χ3 angle, energy value, kcal/mol), Leu92-Glu95 (χ3 angle, −102.37, energy value, 3.86 kcal/mol), and Phe111-Pro117 (χ3 angle,-96.20, energy value,4.14 kcal/mol). These residues have either higher energy level *i.e.* >2 kcal/mol or χ3 angle out of range (<−87 and >+97) were selected on purpose to make them stable. The original and disulfide mutant MEPVC structures are shown in Fig. 5. The primary purpose of *in silico* cloning of the MEPVC was to guide molecular biologists and genetic engineers about the possible cloning sites and predicted levels of expression in a specific expression system for instance here in this study we used *E. coli* K12 system. Prior to cloning, reverse translation of the MEPVC sequence was conducted to have an optimized codon usage as per *E. coli* K12 to yield its max expression. The CAI value of the improved MEPVC sequence is 1 indicating ideal expression of the vaccine [91]. The GC content whereas is 53.2% nearly to the *E. coli* K12 and range within the optimum ranged between 30% and 70%. The cloned MEPVC is shown in Fig. 6.

3.6. In silico immune simulation

Both primary and secondary immune responses seem to play a significant contribution against the pathogen and may be compatible with the actual immune response. The *in silico* host immune system response to the antigen is shown in Fig. 7. High concentration of IgG +IgG and IgM was characterized at the primary response, followed by IgM, IgG1+ IgG2 and IgG1 at both primary and secondary stages with concomitant of antigen reduction. Additionally, robust responses of interleukins and cytokines were observed. All this suggests the efficient immune response and clearance of the pathogen upon subsequent encounters. Elevated B cell population including memory cells and different isotypes in response to the antigen, points to the long lasting formation of memory and isotype switching. The T helper cell population additionally with the cytotoxic T cell and their respective memory development are in strong agreement of strong response to the antigen.

3.7. Interaction of MEPVC and TLR3 promotes formation of strong complex

Bioinformatic modeling driven molecular docking of the designed MEPVC to one representative innate immune response receptor (TLR3) was carried out in order to decode MEPVC potential of binding to the innate immune receptors. This was fundamental to understand as TLR3 is significant in recognition of virus associated molecular patterns and activation of type I interferons and NF-kappa B. The docking assessment predicted top 20 complexes sorted mainly on scoring function along with area size of interacting molecules, actual rigid transformation of complexes and desolvation energy. Following, the complexes were subjected to FireDock web server for refinement assay. This eases in discarding flexibility errors of the docking procedure and provides a deep refinement of the predictions thus limiting the chances of false positive docking calculations. According to the global energy, solution 8 was ranked as top with net global energy of −20.78 kJ/mol (Table 1). This energy is the output of −16.88 kJ/mol attractive van der Waals (vdW), 3.81 kJ/mol repulsive vdW, 8.14 kJ/mol atomic contact energy (ACE), and − 0.93 kJ/mol hydrogen bond energy. The docked conformation and chemical interacting residues of the MEPVC with TLR3 is shown in Fig. 8. Visual inspection of the complex leads to observation of deep binding of the MEPVC at the center of TLR3 and favor rigorously hydrogen and weak van der Waals interactions with various residues of TLR3. Within 3 Å, the MEPVC was noticed to formed interactions with His39,Val55,Asn57,Asp81,Phe84, Val103,Asn105,Gln107,His108,Thr126,Glu127,Ser132,His129,Thr151, His156,Gln174,GLu175,Lys200,Lys201,Glu203,Asn229,Ser256,Asp280, Ser282,Tyr283,Tyr302,Phe304,Tyr307,Lys330,Tyr383,Tyr326,Asn328, His359,Asn361, and Glu363.

3.8. MD simulations supported docked complex stability and equilibrium

The stability of MEPVC with TLR3 was further investigated through MD simulations. The trajectories of MD simulations were used in vital statistical analysis to decode backbone stability and residual flexibility. Root mean square deviation (RMSD) [92,93] was performed first that compute the average distance of backbone carbon alpha atoms of superimposed frames (Fig. 9A). The average RMSD for the system is 3.23 Å with max of 4.4 Å at 24-ns. An initial sudden change in RMSD can be seen up to 2.7-ns that may be due to adjustments adopted by the complex when exposed to dynamics forces and milieu. The second minor RMSD shift can be noticed between 22-ns to 26-ns. Afterward, the system is quite stable with no global and local conformational changes specified. To affirm this observation, the simulation period was extended to 100-ns (S-Fig. 3). As can be seen, the complex is attaining stability as simulation time proceeds and the average RMSD reduced to 3.0 Å. Next, root mean square fluctuations (RMSF) [94] was applied to the system trajectories (Fig. 9B). RMSF is the average residual mobility of complex residues from its mean position. Mean RMSF for the MEPVC-TLR3 complex calculated is 1.60 Å with max of 8.6 Å pointed at the N-terminal of the MEPVC. Most of the interacting residues of the MEPVC with TLR3 are subject to minor fluctuations, a fact in analogy to complex high stability. The thermal residual deviation was assessed afterward by beta-factor (β-factor) [95], the outcomes of which are strongly correlated to the RMSF and hence further affirming system stability (Fig. 9C). The average β-factor of the system analyzed is 88.64 $Å^2$ with max of 1956.23 $Å^2$. Lastly, we evaluated the compactness of the system by means of radius of gyration (Rg) [96] analysis (Fig. 9D). High Rg and low Rg illustrate the magnitude of system compactness and system less tight packing. It further tells us whether the system of interest in order or not. Highly compact system is an indication of system stability and *vice versa*. The mean Rg for our system is 55.8411 Å with max score of 74.884 reflecting higher ordered and compact nature of the system.

3.9. Hydrogen bonds analysis

Hydrogen bonds are dipole-dipole attractive forces and formed when a hydrogen atom bonded to a highly electronegative atom such as F, N, and O is attracted by another electronegative atom [97]. The strength of a hydrogen bond varies from 4 kJ to 50 kJ per mole. Hydrogen bonds are deemed vital in molecular recognition and provide rigidity in achieving stable conformation [98]. The frequency of

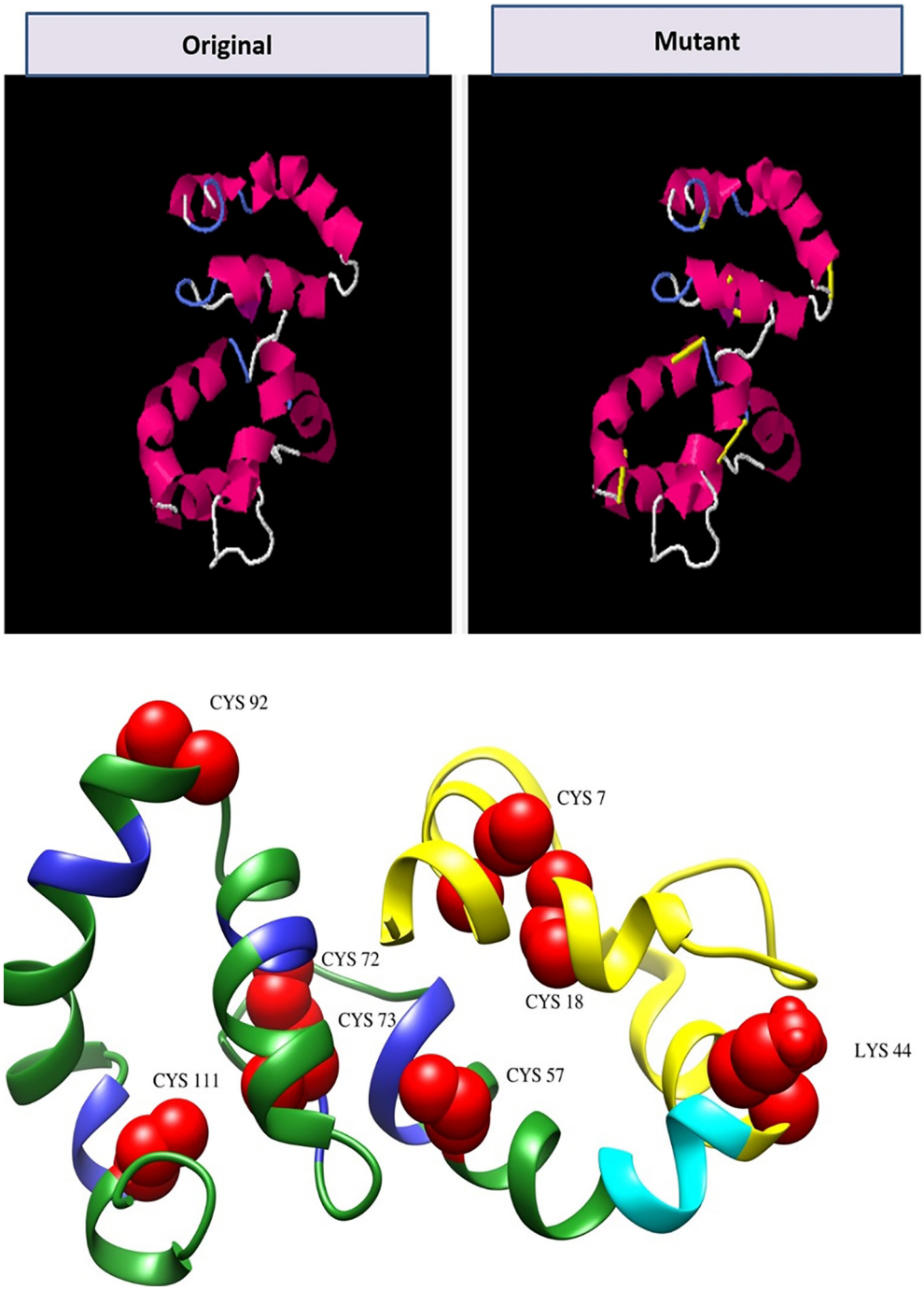

Fig. 5. Disulfide engineering of the MEPVC (Top). The yellow bonds are the disulfide of the mutated residues to increase MEPVC stability. The red spheres in the MEPVC 3D structure (bottom) represent the mutated residues.

hydrogen bonds in each frame of the MD simulation trajectories can be visualized in Fig. 10A. These hydrogen bonds were extracted by means of VMD HBonds plugin and are 104 in number (Table 2). The cut-off distance set is 3.0 Å and cut-off angle 20 degrees. Each residue pair may for one, two or more each of which is counted separately. The min, mean and max number of hydrogen bonds between MEPVC and TLR3 are 1, 5, and 12, respectively.

3.10. Radial distribution function and inter molecular interactions

The distribution and bonding pattern of intermolecular interactions of the MEPVC residues atom(s) with respect to the TLR3 were studied through radial distribution function (RDF) [94,99,100]. RDF mainly describes distance 'r' between two entities and is symbolized by g(r). The factor 'r' is extracted from simulation frames and range from o to ∞ [75]. The hydrogen bonds predicted by VMD were utilized in RDF that shown only 8 interactions between MEPVC and TLR3 with good affinity for each other. In these interactions, TLR3 residues (atoms) are Asp52:OD1, Gln78:HE21, Glu98:OE1, Asp124:OD2, Lys171:HZ3, Asp251:OD2, Glu323:OE2, and Glu328:OE1 are found to have strong radii distribution to their counterpart MEPVC residues (atoms): Arg705:HH11, Arg705: O, Lys679:HZ3, Arg706:HH12, Glu675:OE1, Asn633:HD22, Arg643:HH22, and Tyr638: HH, respectively. The RDF plots for the above said interactions are illustrated in Fig. 10B. The

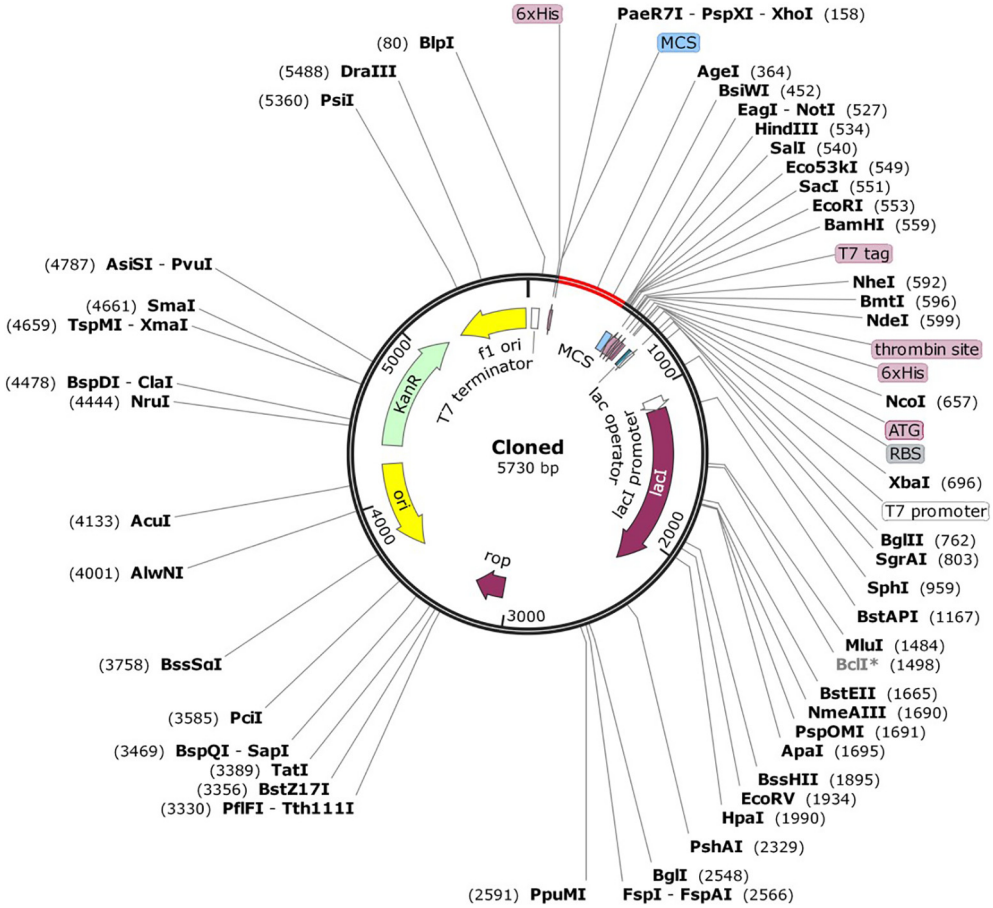

Fig. 6. *In silico* cloning of MEPVC (shown in red) in the pET28a expression vector.

interaction between Asp124-OD2 and Arg706-HH12 has a refined distribution pattern and highest density distribution among all. The max g(r) value for this interaction is 3.51 observed at distance range of 1 Å. This is followed by Glu328-OE1-Tyr638-HH with max g (r) value of 3.26 mostly interacting at distance range of 0.6 Å. The Glu323-OE2-Arg643-HH22 is also much refined having g(r) value of 1.98 and mostly interacts within distance range of 0.6 Å. The remaining interactions density distribution is not confined and vary considerably but important from MEPVC and TLR3 interaction point of view.

3.11. Salt bridges and TLR3-MEPVC stability

Salt bridges are non-covalent in nature and the outcome of interactions between two ionized states [101]. These interactions

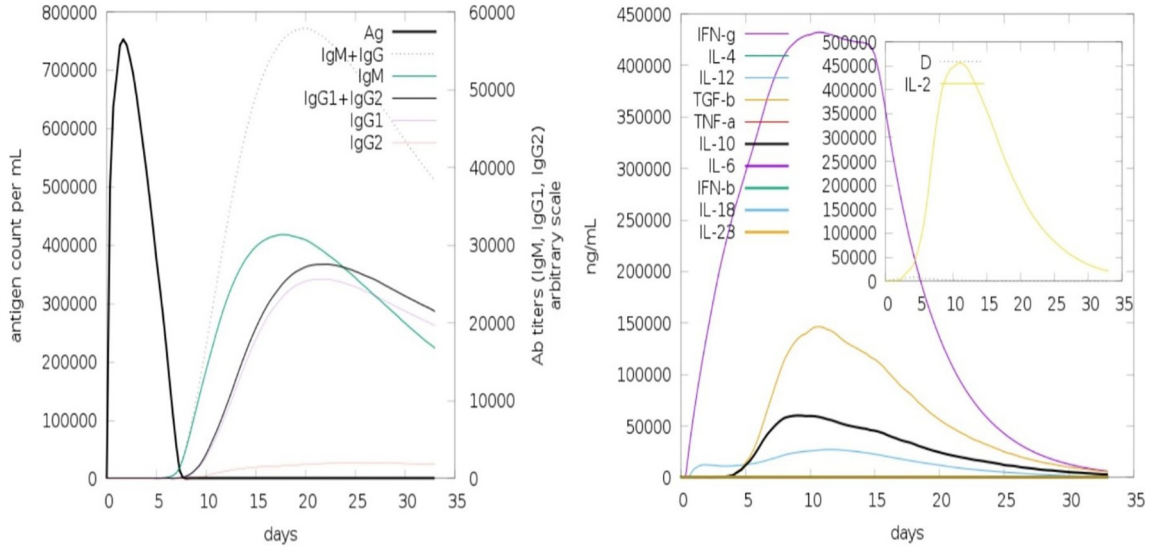

Fig. 7. Computational immune simulation of the host immune system response to the MEPVC. The antibodies titer (left) and cytokines and interleukins concentration (right) to the MEPVC are shown in the figure.

Table 1
Refined PatchDock complexes as an outcome of FireDock assay.

Solution rank	Solution iteration	Docking global energy	Attractive van der Waals energy	Repulsive van der Waals energy	Atomic contact energy	Hydrogen bonding energy
1	8	−20.78	−16.88	3.81	8.14	−0.93
2	3	−9.08	−11.71	1.36	2.37	−0.96
3	4	18.96	−35.55	34.92	8.46	−2.21
4	1	19.29	−12.32	15.65	6.68	−2.85
5	10	23.90	−28.09	18.61	10.32	−2.74
6	7	28.45	−24.64	22.73	12.98	−3.35
7	6	34.28	−29.92	9.59	14.30	−1.52
8	9	41.80	−28.57	10.33	16.83	−4.66
9	5	47.06	−11.90	80.07	4.28	−2.03
10	2	76.58	−40.07	43.84	20.00	−5.59

comprised two parts: an electrostatic interaction and a hydrogen bond. In salt bridges, lysine or arginine typically behave as base where glutamine or aspartate as acid and the bridge is created when carboxylic acid group allows a proton migration to guanidine and amine group in arginine. Salt bridges are the strongest among all non-covalent interactions and contribute to a major extent in biomolecular stability [102–104]. In total, 17 salt bridges were identified between TLR3 and MEPVC within the cut-off distance of 3.2 Å as can be depicted from Fig. 10C. The higher numbers of salt bridges were recorded for TLR3-Glu628 and MEPVC-Lys685. The mean number of salt bridges for this interaction is 18, max, 35, and min,3. The count for other salt bridges from TLR3 to MEPVC is in following order: Asp124-Arg706 (mean, 3, max,7 and min,3), Glu98-Lys679 (mean, 5, max,10 and min,4), Asp402-Arg641 (mean, 12, max, 18 and min,4), Glu146-Arg706 (mean,7 max,11 and min,3), Glu146-Lys679 (mean,5 max,13 and min,2), Glu272-Lys655 (mean,9 max,22 and min,2), Glu323-Arg643 (mean,4 max,8 and min,3), Glu323-Arg646 (mean,5 max,10 and min,3), Glu323-Lys655 (mean,13 max,25 and min,3), Glu328-Arg641 (mean,10 max,17 and min,4), Glu628-Lys699 (mean,9 max,28 and min,2), Glu675-Lys171 (mean,4 max,13 and min,2), Glu675-Lys172 (mean,10 max,16 and min,2), Glu98-Arg705 (mean,9 max,14 and min,5), Glu98-Arg706 (mean, 9 max,14 and min,3), and Glu98-Lys679 (mean, 5 max,11 and min,2).

3.12. Density distribution and local structure movements

The vital hydrogen bond interactions involved between the TLR3 receptor and MEPVC shortlisted by VMD were subjected to a novel AFD analysis to elucidate 3D movements of MEPVC atoms with respect to a reference TLR3 residues atom in simulation time. To this objective, interactions mentioned in the RDF were used in AFD. Preliminary investigation suggested that only three interactions: TLR3-Asp52-MEPVC-Arg705, TLR3-Glu328-MEPVC-Tyr638, and TLR3-Glu323-MEPVC-Arg643 are mainly represented frequently and found in most of the simulation frames. The TLR3-Asp52-MEPVC-Arg705 is uncovered in 4997 frames, TLR3-Glu328-MEPVC-Tyr638 in 4988, and TLR3-Glu323-MEPVC-Arg643 in 4985 making these interactions ideal for interpreting density distribution of the interactions on XYZ planes and also appropriate for gaining ideas about conformational changes of the interacting atoms with respect to each other. As the local structure movements and rotations are responsible for functional shifts, their understanding in our system is important to be unveiled. For TLR3-Asp52-MEPVC-Arg705 (Fig. 11A), the density distribution is not uniform, dispersed and behave flexibility in affinity on all three axis for the receptor atom. Parallel, the strength of interaction is also observed affected due to these minor structural movements of the MEPVC residue atom. Though, the mentioned interaction depicts

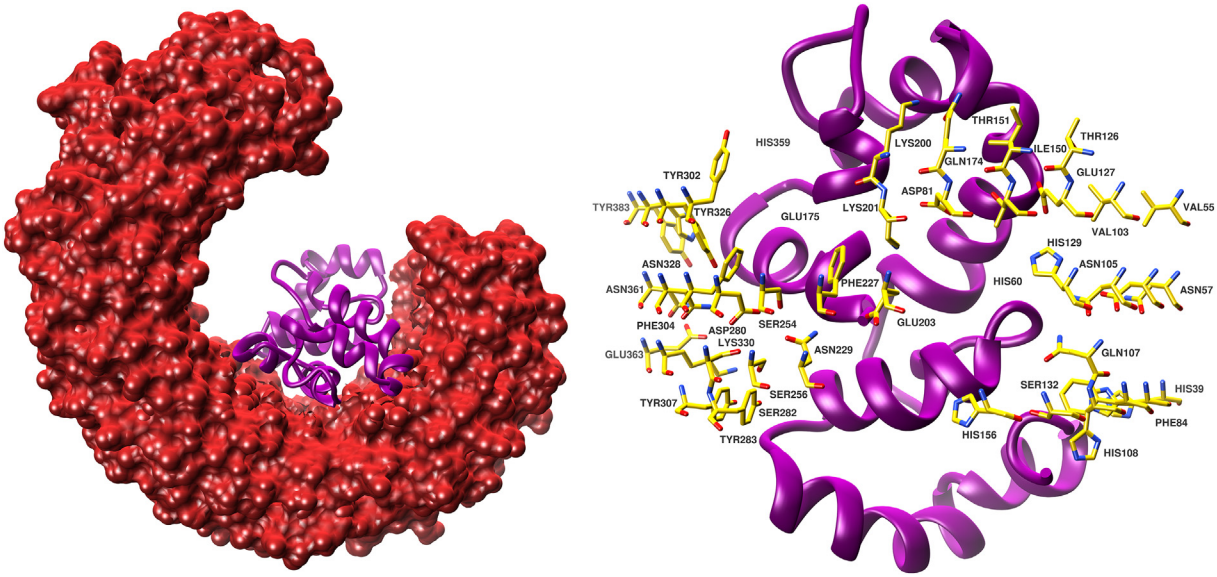

Fig. 8. A. MEPVC (shown in dark magenta) conformation with respect to the TLR3 receptor (shown in firebrick surface). B. TLR3 interacting residues (shown in yellow sticks) are involved in hydrophobic and hydrophilic interactions with the MEPVC.

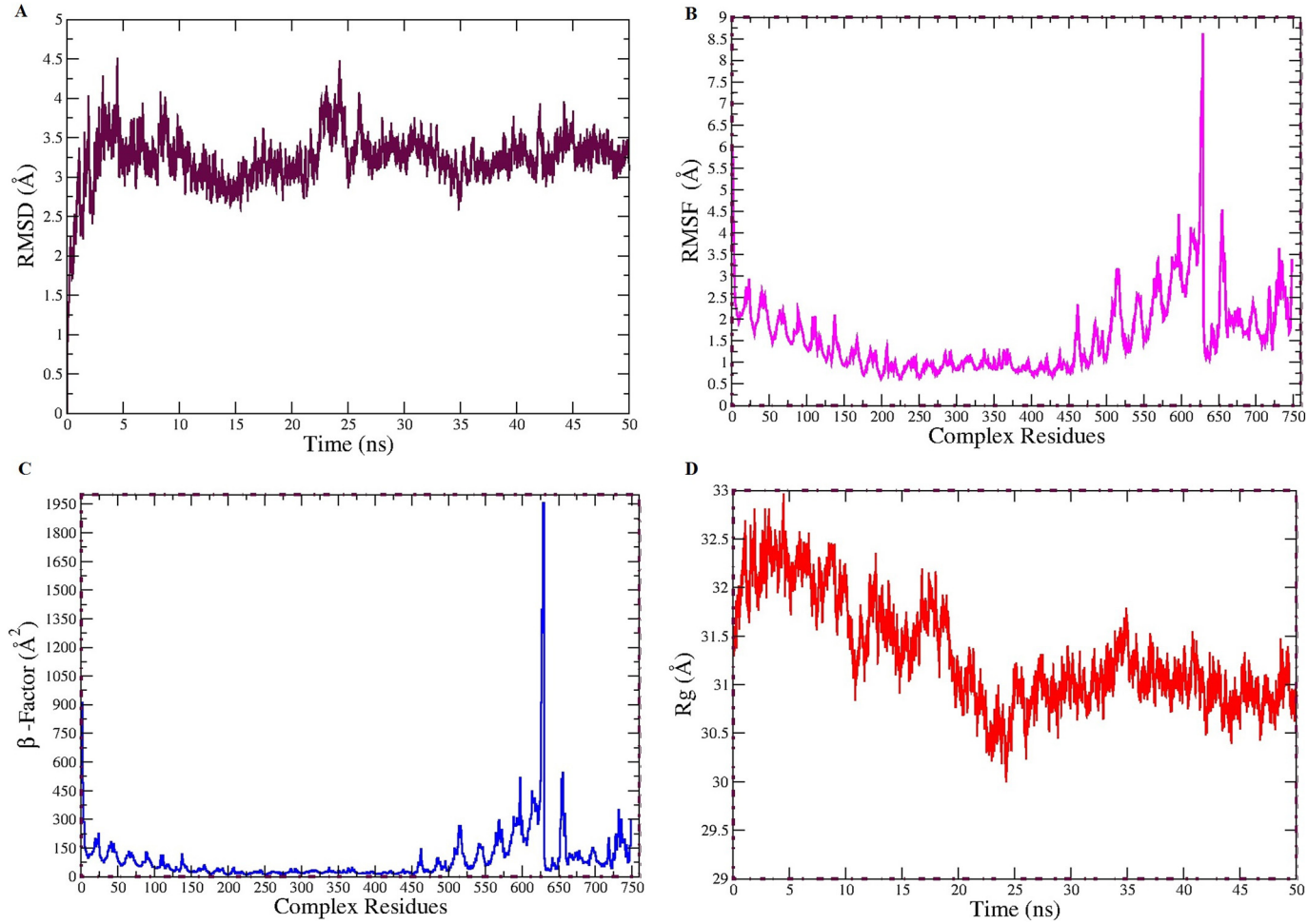

Fig. 9. MD simulations based analysis of RMSD (A), RMSF (B), β-factor (C), and Rg (D).

MEPVC is still within the vicinity of the TLR3 reference residue and enjoys this interaction flexibility with the said MEPVC residue during simulation. TLR3-Glu328-MEPVC-Tyr638 interaction (Fig. 11B) has less distribution area and has much higher intensity illustrating strong affinity of the interacting atoms for each other. It also gives an idea of the lesser movements of the atoms with respect to each other, an indication of a correct system conformation. The distribution area TLR3-Glu323-MEPVC-Arg643 is much dispersed though high intensity of the interaction can be seen in close vicinity (Fig. 11C).

3.13. MEPVC-TLR3 complex revealed high binding free energies

The ΔTOTAL free energy in both GB and PB models are revealed favorable MEPVC-TLR3 complex in solvation. The net GB and PB contribution for the MEPVC-TLR3 system is −53.81 kcal/mol and − 89.02 kcal/mol, respectively. To this energy, a high contribution was noticed from gas phase energy (ΔG gas) compared to highly insignificant contributions from solvation energy (ΔG solv). In GB model, the ΔG gas energy for the system is −1889.76 kcal/mol whereas in PB model this energy is −1889.76 kcal/mol. The ΔG solv energy in case of GB is 1835.95 kcal/mol while in case of PB it is 1800.74 kcal/mol. The electrostatic contribution to the system estimated by MM force field in both methods is highly favorable to the net energy (*i.e.* −1752.51 kcal/mol). Likewise, the van der Waals contribution form MM is also reported significant in system stability (−137.25 kcal/mol). The electrostatic contribution (EGB and EPB) to the ΔG solv is found prime parameter leading to non-favorable contribution of ΔG solv energy in both techniques. The surface area energy (ESURF) computed in GB method is −20.22 kca/mol. In PB, ENPOLAR and EDISPER are the repulsive and attractive free energy and are −17.44 kcal/mol and 0 kcal/mol, respectively. Individual binding free energy for MEPVC-TLR3 complex, TLR3 receptor, MEPVC and net energy is tabulated in Table 3.

3.14. Net energy decomposition discovered hot-spot residues

The net free energy of the simulated system was subjected to per residues and pairwise residues decomposition to point residues that contribute majorly in system stabilization and lower energy. Molecular docking simulation studies demonstrated 64 residues from the TLR3 receptor that are in direct contact with the MEPVC but per residue decomposition assay illustrated that among the residues only Hie31, Phe55, Glu98, Hie100, Met102, Ile121, Thr122, Asp124, Glu146, Glu146, Glu174, Ser176, Phe198, Asn200, Ser225, Met249, Asp251, Tyr254, Tyr273, Phe275, Tyr278, Tyr297, Glu323, Hie324, and Tyr348 are hotspot as they contribute rigorously in binding interaction with MEPVC at the docked site. The side chain of Hie100, Thr122, Asp124, Glu174, Ser176, Tyr254, and Glu323 contribute significantly in chemical interactions and have energy value in following order: −2.86602 kcal/mol, −3.71782 kcal/mol, −3.77019 kcal/mol, −3.80724 kcal/mol, −2.71475 kcal/mol, −2.40187 kcal/mol, and − 3.54158 kcal/mol, respectively. To these TLR3 hotspot residues, the MEPVC interacting residues were also observed in quite lower energies illustrating high affinity for the receptor residues for chemical interactions. From

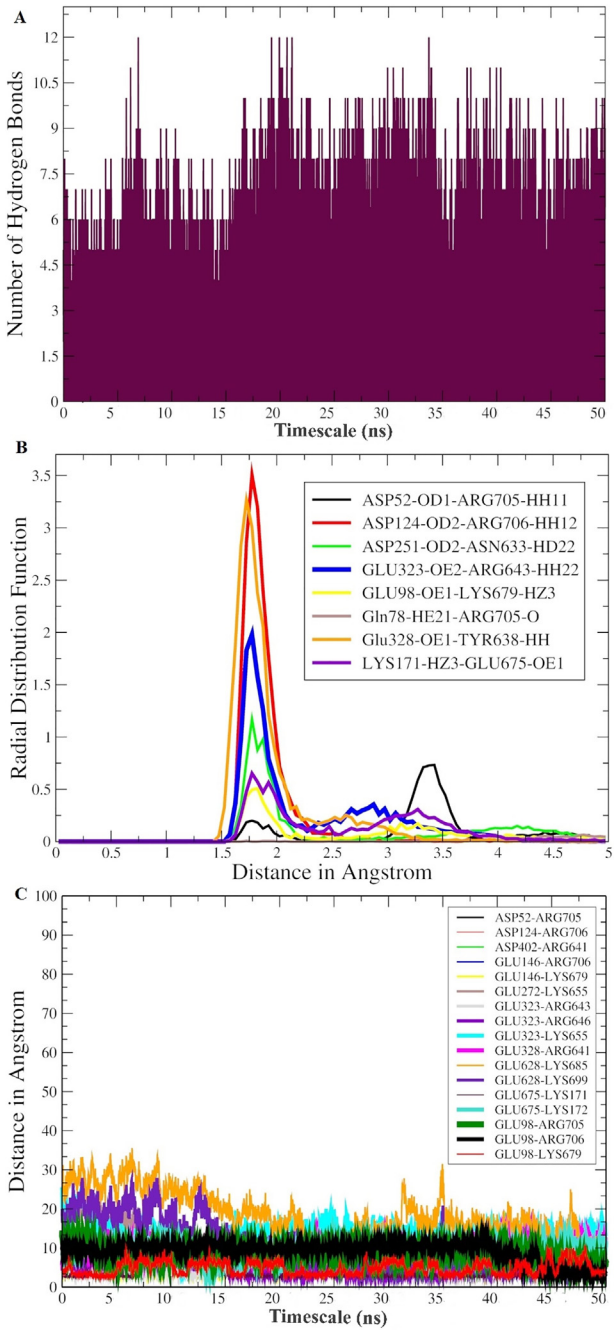

Fig. 10. Number of hydrogen bonds in each frame of MD simulation trajectories (A), RDF plots for hydrogen bonds in TLR3-MEPVC critical in interaction and stability (B), and salt bridges formed between TLR3 and MEPVC (C) during simulation time.

MEPVC, ASN633 (−4.11274 kcal/mol), TYR638 (2.83056 kcal/mol), VAL642 (−2.12862 kcal/mol), ARG643 (−6.59981 kcal/mol), GLU675 (−2.57531 kcal/mol), TRP682 (−2.54238 kcal/mol), ARG705 (−3.1088 kcal/mol), and ARG706 (−4.45088 kcal/mol) are favorable residues in stable complex formation. The hotspot residues for both TLR3 and MEPVC are shown in S-Fig. 4.

3.15. Frame –wise energy decomposition

The binding free energy of the TLR3-MEPVC complex, TLR3 receptor, MEPVC, and the net system energy is further decomposed into 100 frames extracted from simulation trajectories (S-Fig. 5). This information deemed vital in predicting the simulation time where higher intermolecular affinity was observed and can guide about the

Table 2
Hydrogen bonds between TLR3 and MEPVC formed during MD simulations. Only hydrogen bonds with occupancy >5% are shown in the table.

Donor	Acceptor	Occupancy
Arg705-Side	Asp52-Side	98.52%
Arg643-Side	Glu323-Side	83.46%
Arg706-Side	Asp124-Side	36.24%
Tyr638-Side	Glu328-Side	34.02%
Lys171-Side	Glu675-Side	29.48%
Arg646-Side	Glu323-Side	25.80%
Thr697-Side	Thr629-Main	23.12%
Gln722-Main	Tyr254-Side	22.94%
Lys679-Side	Glu98-Side	14.34%
Asn633-Side	Asp251-Side	14.20%
Arg706-Side	Glu98-Side	12.14%
Lys699-Side	Thr629-Side	11.72%
Arg643-Side	Hie324-Side	11.54%
Lys679-Side	Thr122-Side	8.52%
Lys679-Side	Glu146-Side	7.34%
Lys699-Side	Glu628-Side	6.90%
Tyr278-Side	Pro720-Main	6.86%
Gly630-Main	Glu174-Side	6.26%
Lys699-Side	Thr629-Main	5.64%
Arg706-Side	Glu146-Side	5.36%

most suitable docked conformation. In general, the complex, receptor and construct energies are higher in PB compared to GB but for the total energy, the PB energies are quite lower for frames in contrast to GB. For the complex, the min, max and average binding energy reported are −67,264.7 kcal/mol, −66,901.5 kcal/mol, and − 67,069.5 kcal/mol, respectively in GB. The PB max frame energy is −66,751.4 kcal/mol, min is −67,120.2 kcal/mol and average is −66,921.6 kcal/mol. The GB receptor max is −57,111.5 kcal/mol whereas the min is −57,381.8 kcal/mol, and average is −57,242 kcal/mol. For PB receptor, max of −56,961 kcal/mol of energy was noticed compared to the min of −57,249 kcal/mol. The MEPVC is showing high average energy (−9773.16 kcal/mol in GB and − 9742.47 kcal/mol in PB), max (−9672.42 kcal/mol in GB and − 9644.54 kcal/mol in PB) and min (−9838.37 kcal/mol in GB and − 9811.3 kcal/mol in PB). The net average energy for frames in GB is −53.8024 kcal/mol (max, −33.978 kcal/mol and min, −72.791 kcal/mol) and in PB is −89.0949 kcal/mol (max, −65.151 kcal/mol and − 120.384 kcal/mol).

3.16. Pair-wise energy contribution

Pair-wise energy contribution to the system total energy was accomplished in order to understand pair residues roles from both TLR3 and MEPVC in complex stability. We found that the Thr122 and Asp124 (−4.56 kcal/mol in GB and − 5.45 kcal/mol in PB), Glu174 and Ser176 (−3.45 kcal/mol in GB and − 3.77 kcal/mol in PB), Glu323 and Hie324 (−2.86 kcal/mol in GB and − 3.99 kcal/mol in PB) of TLR3 receptor have high combined contribution to the total energy. In case of MEPVC, Asn633 and Thr634 (−3.21 kcal/mol in both GB and PB), Val642 and Arg643 (−5,87 kcal/mol in GB and − 3.27 kcal/mol in PB) and Arg705 and Arg708 (−2.74 kcal/mol and 2.04 kcal/mol).

3.17. Binding entropy calculation

Detail findings of entropy calculation are tabulated in Table 4. The net value of −52.52 kcal/mol symbolizes the fact that microstate level of the system decreases when the TLR3 receptor and MEPVC bind to make complex. The decrease in the available microstate level also points towards MEPVC efficient trapping at the docked site of TLR3.

4. Conclusions

Taken together, we characterized SARS-CoV-2 spike glycoprotein for antigenic peptides and proposed a MEPVC by means of several

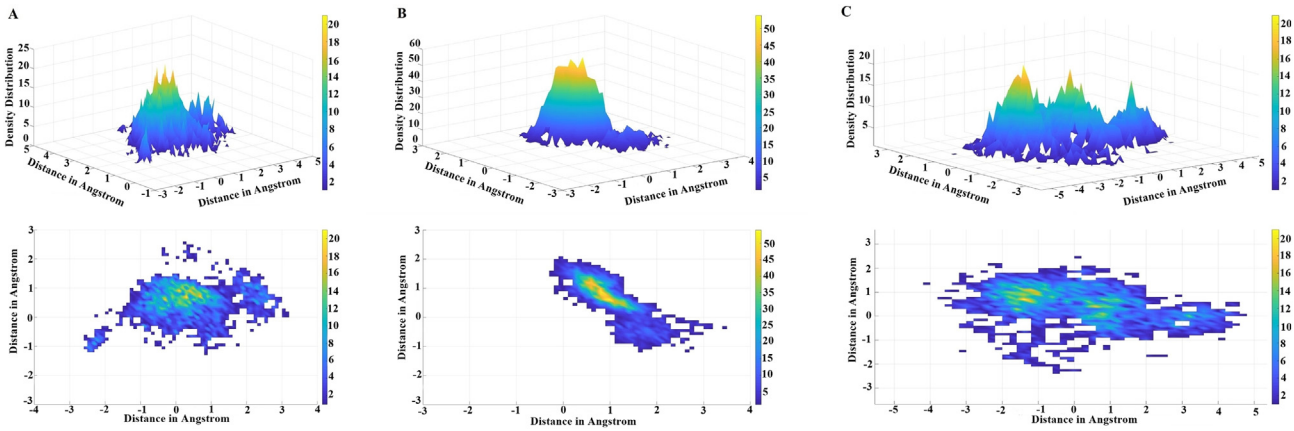

Fig. 11. AFD plots for TLR3-Asp52-MEPVC-Arg705 interaction (A), AFD plots for TLR3-Glu328-MEPVC-Tyr638 (B), AFD plot for TLR3-Glu323-MEPVC-Arg643 (C). Three points perspective AFD plots are placed at the top in the figure for the interactions whereas top down views are placed at the bottom.

Table 3
Binding free energies for the MEPVC-TLR3 system.

GB				PB			
Differences (MEPVC-TLR3 - TLR3-vaccine MEPVC)				Differences (MEPVC-TLR3 - TLR3-vaccine MEPVC)			
Energy component	Average	Std. dev.	Std. err. of mean	Energy component	Average	Std. dev.	Std. err. of mean
VDWAALS	−137.25	7.9	0.79	VDWAALS	−137.25	7.9	0.79
EEL	−1752.51	48.84	4.88	EEL	−1752.51	48.84	4.88
EGB	1856.18	45.486	4.54	EPB	1818.19	45.33	4.53
ESURF	−20.22	1.29	0.12	ENPOLAR	−17.44	0.51	0.05
–	–	–	–	EDISPER	0	0	0
ΔG gas	−1889.76	48.37	4.83	ΔG gas	−1889.76	48.37	4.83
ΔG solv	1835.95	45.19	4.51	ΔG solv	1800.74	45.2	4.52
ΔTotal	−53.81	8.1955	0.8195	Δ Total	−89.02	11.55	1.15

computational immunological methods and biophysical calculations. The outcomes of this study could save time and associated costs that go into experimental epitope targets study. The MEPVC is capable of activating all components of the host immune system, have suitable structural and physicochemical properties. Also, it seems to have very stable dynamics with TLR3 innate immune receptor and thus has higher chances of presentation to the host immune system. However, additional *in vivo* and *in vitro* experimentations are needed to disclose its real potential in the fight against COVID-19.

Table 4
Entropy results for the TLR3-MEPVC system. All values are in kcal/mol.

	Entropy term	Average	Std. dev.
Complex	Translational	17.80	0.00
	Rotational	17.89	0.00
	Vibrational	2792.19	2.33
	Total	2827.88	2.33
TLR3	Translational	16.81	0.00
	Rotational	17.56	0.00
	Vibrational	2786.86	3.03
	Total	2821.23	3.03
Vaccine MEPVC	Translational	14.01	0.00
	Rotational	11.59	0.04
	Vibrational	33.57	0.04
	Total	59.17	0.00
ΔS total	DELTA S total	−52.52+/−	3.30

CRediT authorship contribution statement

Saba Ismail: Methdology, Writing - original draft, Data curation, Formal Analsysis. **Sajjad Ahmad:** Data curation, Writing - original draft, Visualization. **Syed Sikander Azam:** Supervision, Conceptualization, Methodology, Software, Resources, Funding acquisition, Writing - review & editing, Project administration.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Acknowledgment

Authors are highly grateful to the Pakistan-United States Science and Technology Cooperation Program (Grant No. Pak-US/2017/360) for granting the financial assistance.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.molliq.2020.113612.

References

[1] K. Dhama, K. Sharun, R. Tiwari, M. Dadar, Y.S. Malik, K.P. Singh, W. Chaicumpa, COVID-19, an emerging coronavirus infection: advances and prospects in designing and developing vaccines, immunotherapeutics, and therapeutics, Hum. Vaccin. Immunother. (2020) 1–7.
[2] L. Wang, Y. Wang, D. Ye, Q. Liu, A review of the 2019 Novel coronavirus (COVID-19) based on current evidence, Int. J. Antimicrob. Agents 55 (2020) 105948.
[3] Y. Chen, Q. Liu, D. Guo, Emerging coronaviruses: genome structure, replication, and pathogenesis, J. Med. Virol. 92 (2020) 418–423.
[4] Y. Yang, F. Peng, R. Wang, K. Guan, T. Jiang, G. Xu, J. Sun, C. Chang, The deadly coronaviruses: the 2003 SARS pandemic and the 2020 novel coronavirus epidemic in China, J. Autoimmun. 111 (2020) 102434.
[5] Z.-W. Ye, S. Yuan, K.-S. Yuen, S.-Y. Fung, C.-P. Chan, D.-Y. Jin, Zoonotic origins of human coronaviruses, Int. J. Biol. Sci. 16 (2020) 1686–1697.
[6] J.T. Wu, K. Leung, M. Bushman, N. Kishore, R. Niehus, P.M. de Salazar, B.J. Cowling, M. Lipsitch, G.M. Leung, Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China, Nat. Med. (2020) 1–5.
[7] C.P.E.R.E. Novel, et al., The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China, Zhonghua Liu Xing Bing Xue Za Zhi= Zhonghua Liuxingbingxue Zazhi 41 (2020) 145.
[8] A.C. Walls, Y.-J. Park, M.A. Tortorici, A. Wall, A.T. McGuire, D. Veesler, Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein, Cell 181 (2020) 281–292.e6.
[9] F. Li, Structure, function, and evolution of coronavirus spike proteins, Annu. Rev. Virol. 3 (2016) 237–261.

[10] B.J. Bosch, R. van der Zee, C.A.M. de Haan, P.J.M. Rottier, The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex, J. Virol. 77 (2003) 8801–8811.
[11] D. Wrapp, N. Wang, K.S. Corbett, J.A. Goldsmith, C.-L. Hsieh, O. Abiona, B.S. Graham, J.S. McLellan, Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation, Science (80-.). 367 (2020) 1260–1263.
[12] A.C. Walls, M.A. Tortorici, J. Snijder, X. Xiong, B.-J. Bosch, F.A. Rey, D. Veesler, Tectonic conformational changes of a coronavirus spike glycoprotein promote membrane fusion, Proc. Natl. Acad. Sci. 114 (2017) 11157–11162.
[13] M. Gui, W. Song, H. Zhou, J. Xu, S. Chen, Y. Xiang, X. Wang, Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding, Cell Res. 27 (2017) 119–129.
[14] J. Pallesen, N. Wang, K.S. Corbett, D. Wrapp, R.N. Kirchdoerfer, H.L. Turner, C.A. Cottrell, M.M. Becker, L. Wang, W. Shi, et al., Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen, Proc. Natl. Acad. Sci. 114 (2017) E7348–E7357.
[15] A.C. Walls, X. Xiong, Y.-J. Park, M.A. Tortorici, J. Snijder, J. Quispe, E. Cameroni, R. Gopal, M. Dai, A. Lanzavecchia, Others, unexpected receptor functional mimicry elucidates activation of coronavirus fusion, Cell 176 (2019) 1026–1039.
[16] Y. Yuan, D. Cao, Y. Zhang, J. Ma, J. Qi, Q. Wang, G. Lu, Y. Wu, J. Yan, Y. Shi, Others, Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains, Nat, Commun 8 (2017) 15092.
[17] J. Goo, Y. Jeong, Y.-S. Park, E. Yang, D.-I. Jung, S. Rho, U. Park, H. Sung, P.-G. Park, J. Choi, et al., Characterization of novel monoclonal antibodies against MERS-coronavirus spike protein, Virus Res. (2020) 197863.
[18] R. Chen, J. Fu, J. Hu, C. Li, Y. Zhao, H. Qu, X. Wen, S. Cao, Y. Wen, R. Wu, et al., Identification of the immunodominant neutralizing regions in the spike glycoprotein of porcine deltacoronavirus, Virus Res. 276 (2020), 197834. .
[19] M. Bhattacharya, A.R. Sharma, P. Patra, P. Ghosh, G. Sharma, B.C. Patra, S.-S. Lee, C. Chakraborty, Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): Immunoinformatics approach, J. Med. Virol. 19 (2020) 149–150.
[20] G. Li, E. De Clercq, Therapeutic Options for the 2019 Novel Coronavirus (2019-nCoV), 2020.
[21] Y. Wan, J. Shang, R. Graham, R.S. Baric, F. Li, Receptor recognition by the novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS coronavirus, J. Virol. 94 (2020).
[22] W. Li, M.J. Moore, N. Vasilieva, J. Sui, S.K. Wong, M.A. Berne, M. Somasundaran, J.L. Sullivan, K. Luzuriaga, T.C. Greenough, Others, Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus, Nature 426 (2003) 450–454.
[23] M. Hoffmann, H. Kleine-Weber, N. Krüger, M.A. Mueller, C. Drosten, S. Pöhlmann, The novel coronavirus 2019 (2019-nCoV) uses the SARS-coronavirus receptor ACE2 and the cellular protease TMPRSS2 for entry into target cells, BioRxiv (2020).
[24] R.N. Kirchdoerfer, N. Wang, J. Pallesen, D. Wrapp, H.L. Turner, C.A. Cottrell, K.S. Corbett, B.S. Graham, J.S. McLellan, A.B. Ward, Stabilized coronavirus spikes are resistant to conformational changes induced by receptor recognition or proteolysis, Sci. Rep. 8 (2018) 1–11.
[25] J.F.-W. Chan, S. Yuan, K.-H. Kok, K.K.-W. To, H. Chu, J. Yang, F. Xing, J. Liu, C.C.-Y. Yip, R.W.-S. Poon, Others, A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster, Lancet 395 (2020) 514–523.
[26] G. Zhou, Q. Zhao, Perspectives on therapeutic neutralizing antibodies against the Novel coronavirus SARS-CoV-2, Int. J. Biol. Sci. 16 (2020) 1718–1723.
[27] P.C.Y. Woo, S.K.P. Lau, B.H.L. Wong, K. Chan, C. Chu, H. Tsoi, Y. Huang, J.S.M. Peiris, K. Yuen, Longitudinal profile of immunoglobulin G (IgG), IgM, and IgA antibodies against the severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein in patients with pneumonia due to the SARS coronavirus, Clin. Diagn. Lab. Immunol. 11 (2004) 665–668.
[28] W.-C. Cao, W. Liu, P.-H. Zhang, F. Zhang, J.H. Richardus, Disappearance of antibodies to SARS-associated coronavirus after recovery, N. Engl. J. Med. 357 (2007) 1162–1163.
[29] E. Traggiai, S. Becker, K. Subbarao, L. Kolesnikova, Y. Uematsu, M.R. Gismondo, B.R. Murphy, R. Rappuoli, A. Lanzavecchia, An efficient method to make human monoclonal antibodies from memory B cells: potent neutralization of SARS coronavirus, Nat. Med. 10 (2004) 871–875.
[30] L. Liu, Q. Wei, Q. Lin, J. Fang, H. Wang, H. Kwok, H. Tang, K. Nishiura, J. Peng, Z. Tan, et al., Anti–spike IgG causes severe acute lung injury by skewing macrophage responses during acute SARS-CoV infection, JCI Insight 4 (2019).
[31] W. Li, M. Joshi, S. Singhania, K. Ramsey, A. Murthy, Peptide vaccine: progress and challenges, Vaccines 2 (2014) 515–536.
[32] L. Zhang, Multi-epitope vaccines: a promising strategy against tumors and viral infections, Cell. Mol. Immunol. 15 (2018) 182.
[33] H. Sugiarto, P.-L. Yu, Avian antimicrobial peptides: the defense role of $β$-defensins, Biochem. Biophys. Res. Commun. 323 (2004) 721–727.
[34] A. Mittal, S. Sasidharan, S. Raj, S.N. Balaji, P. Saudagar, Exploring the Zika genome to design a potential multiepitope vaccine using an immunoinformatics approach, Int. J. Pept. Res. Ther. (2020) 1–10.
[35] A. Ikram, T. Zaheer, F.M. Awan, A. Obaid, A. Naz, R. Hanif, R.Z. Paracha, A. Ali, A.K. Naveed, H.A. Janjua, Exploring NS3/4A, NS5A and NS5B proteins to design conserved subunit multi-epitope vaccine against HCV utilizing immunoinformatics approaches, Sci. Rep. 8 (2018), 16107. .
[36] M. Karplus, J.A. McCammon, Molecular dynamics simulations of biomolecules, Nat. Struct. Mol. Biol. 9 (2002) 646.
[37] S. Genheden, U. Ryde, The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities, Expert Opin. Drug Discov. 10 (2015) 449–461, https://doi.org/10.1517/17460441.2015.1032936.
[38] R. Vita, S. Mahajan, J.A. Overton, S.K. Dhanda, S. Martini, J.R. Cantrell, D.K. Wheeler, A. Sette, B. Peters, The immune epitope database (IEDB): 2018 update, Nucleic Acids Res. 47 (2018) D339–D343.
[39] M.C. Jespersen, B. Peters, M. Nielsen, P. Marcatili, BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes, Nucleic Acids Res. 45 (2017) W24–W29.
[40] M. Andreatta, M. Nielsen, Gapped sequence alignment using artificial neural networks: application to the MHC class I system, Bioinformatics 32 (2016) 511–517.
[41] P. Wang, J. Sidney, Y. Kim, A. Sette, O. Lund, M. Nielsen, B. Peters, Peptide binding predictions for HLA DR, DP and DQ molecules, BMC Bioinformatics 11 (2010) 568.
[42] P. Guan, I.A. Doytchinova, C. Zygouri, D.R. Flower, MHCPred: a server for quantitative prediction of peptide–MHC binding, Nucleic Acids Res. 31 (2003) 3621–3624.
[43] A. Hassan, A. Naz, A. Obaid, R.Z. Paracha, K. Naz, F.M. Awan, S.A. Muhmmad, H.A. Janjua, J. Ahmad, A. Ali, Pangenome and immuno-proteomics analysis of Acinetobacter baumannii strains revealed the core peptide vaccine targets, BMC Genomics 17 (2016) 732.
[44] A. Garg, D. Gupta, VirulentPred: a SVM based prediction method for virulent proteins in bacterial pathogens, BMC Bioinformatics 9 (2008) 62.
[45] I.A. Doytchinova, D.R. Flower, VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines, BMC Bioinformatics 8 (2007) 4.
[46] I. Dimitrov, D.R. Flower, I. Doytchinova, AllerTOP-a server for in silico prediction of allergens, BMC Bioinformatics (2013) S4.
[47] S. Gupta, P. Kapoor, K. Chaudhary, A. Gautam, R. Kumar, G.P.S. Raghava, Peptide toxicity prediction, Comput. Pept, Springer 2015, pp. 143–157.
[48] S.K. Dhanda, P. Vir, G.P.S. Raghava, Designing of interferon-gamma inducing MHC class-II binders, Biol. Direct 8 (2013) 30.
[49] H.-H. Bui, J. Sidney, W. Li, N. Fusseder, A. Sette, Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines, BMC Bioinformatics 8 (2007) 361.
[50] M. Ali, R.K. Pandey, N. Khatoon, A. Narula, A. Mishra, V.K. Prajapati, Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection, Sci. Rep. 7 (2017) 9232.
[51] E. ProtParam, ExPASy-ProtParam Tool, 2017.
[52] J. Cheng, A.Z. Randall, M.J. Sweredoski, P. Baldi, SCRATCH: a protein structure and structural feature prediction server, Nucleic Acids Res. 33 (2005) W72–W76.
[53] B. Giardine, C. Riemer, R.C. Hardison, R. Burhans, L. Elnitski, P. Shah, Y. Zhang, D. Blankenberg, I. Albert, J. Taylor, et al., Galaxy: a platform for interactive large-scale genome analysis, Genome Res. 15 (2005) 1451–1455.
[54] L. Heo, H. Park, C. Seok, GalaxyRefine: protein structure refinement driven by side-chain repacking, Nucleic Acids Res. 41 (2013) W384–W388.
[55] D.B. Craig, A.A. Dombkowski, Disulfide by design 2.0: a web-based tool for disulfide engineering in proteins, BMC Bioinformatics 14 (2013) 346.
[56] E. Angov, Codon usage: nature's roadmap to expression and folding of proteins, Biotechnol. J. 6 (2011) 650–659.
[57] A. Grote, K. Hiller, M. Scheer, R. Münch, B. Nörtemann, D.C. Hempel, D. Jahn, JCat: a novel tool to adapt codon usage of a target gene to its potential expression host, Nucleic Acids Res. 33 (2005) W526–W531.
[58] S.W. Brubaker, K.S. Bonham, I. Zanoni, J.C. Kagan, Innate immune pattern recognition: a cell biological perspective, Annu. Rev. Immunol. 33 (2015) 257–290.
[59] N. Rapin, O. Lund, F. Castiglione, C-Immsim 10.1 server, 2012.
[60] G.M. Morris, M. Lim-Wilby, Molecular docking, Mol. Model. Proteins, Springer 2008, pp. 365–382.
[61] F.L. Rock, G. Hardiman, J.C. Timans, R.A. Kastelein, J.F. Bazan, A family of human receptors structurally related to Drosophila toll, Proc. Natl. Acad. Sci. 95 (1998) 588–593.
[62] T. Kawai, S. Akira, Signaling to NF-$κ$B by toll-like receptors, Trends Mol. Med. 13 (2007) 460–469.
[63] D. Greenwood, R.C.B. Slack, M.R. Barer, W.L. Irving, Medical Microbiology E-Book: A Guide to Microbial Infections: Pathogenesis, Immunity, Laboratory Diagnosis and Control. With STUDENT CONSULT Online Access, Elsevier Health Sciences, 2012.
[64] L. Alexopoulou, A.C. Holt, R. Medzhitov, R.A. Flavell, Recognition of double-stranded RNA and activation of NF-$κ$B by toll-like receptor 3, Nature 413 (2001) 732–738.
[65] D. Schneidman-Duhovny, Y. Inbar, R. Nussinov, H.J. Wolfson, PatchDock and SymmDock: servers for rigid and symmetric docking, Nucleic Acids Res. 33 (2005) W363–W367.
[66] N. Andrusier, R. Nussinov, H.J. Wolfson, FireDock: fast interaction refinement in molecular docking, Proteins Struct. Funct. Bioinforma. 69 (2007) 139–159.
[67] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF chimera—a visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612.
[68] S. Ahmad, K.E. Ranaghan, S.S. Azam, Combating tigecycline resistant Acinetobacter baumannii: a leap forward towards multi-epitope based vaccine discovery, Eur. J. Pharm. Sci. 132 (2019) 1–17.
[69] D.A. Case, D.S. Cerutti, T.E. Cheateham, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, D. Greene, N. Homeyer, et al., AMBER16 Package, Univ. California, San Fr, 2016.
[70] D.A. Case, V. Babin, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham III, T.A. Darden, R.E. Duke, H. Gohlke, et al., The FF14SB force field, Amber 14 (2014) 29–31.
[71] J.A. Izaguirre, D.P. Catarello, J.M. Wozniak, R.D. Skeel, Langevin stabilization of molecular dynamics, J. Chem. Phys. 114 (2001) 2090–2098.

[72] V. Kräutler, W.F. Van Gunsteren, P.H. Hünenberger, A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations, J. Comput. Chem. 22 (2001) 501–508.
[73] D.R. Roe, T.E. Cheatham III, PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data, J. Chem. Theory Comput. 9 (2013) 3084–3095.
[74] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38.
[75] S. Raza, S.S. Azam, AFD: an application for bi-molecular interaction using axial frequency distribution, J. Mol. Model. 24 (2018)https://doi.org/10.1007/s00894-018-3601-3.
[76] B.R. Miller, T.D. McGee, J.M. Swails, N. Homeyer, H. Gohlke, A.E. Roitberg, MMPBSA.py: an efficient program for end-state free energy calculations, J. Chem. Theory Comput. 8 (2012) 3314–3321, https://doi.org/10.1021/ct300418h.
[77] T. Hou, J. Wang, Y. Li, W. Wang, Assessing the Performance of the MM_PBSA and MM_GBSA Methods. 1. The Accuracy.pdf, 2011 69–82.
[78] A. Naz, F.M. Awan, A. Obaid, S.A. Muhammad, R.Z. Paracha, J. Ahmad, A. Ali, Identification of putative vaccine candidates against Helicobacter pylori exploiting exoproteome and secretome: a reverse vaccinology based approach, Infect. Genet. Evol. 32 (2015) 280–291.
[79] Y. He, Z. Xiang, H.L.T. Mobley, Vaxign: the first web-based vaccine design program for reverse vaccinology and applications for vaccine development, Biomed. Res. Int. 2010 (2010).
[80] K. Naz, A. Naz, S.T. Ashraf, M. Rizwan, J. Ahmad, J. Baumbach, A. Ali, PanRV: Pangenome-reverse vaccinology approach for identifications of potential vaccine candidates in microbial pangenome, BMC Bioinformatics 20 (2019) 123.
[81] V. Solanki, M. Tiwari, V. Tiwari, Prioritization of potential vaccine targets using comparative proteomics and designing of the chimeric multi-epitope vaccine against Pseudomonas aeruginosa, Sci. Rep. 9 (2019).
[82] T.M. Wizemann, J.E. Adamou, S. Langermann, Adhesins as targets for vaccine development, Emerg. Infect. Dis. 5 (1999) 395.
[83] D. Barh, N. Barve, K. Gupta, S. Chandra, N. Jain, S. Tiwari, N. Leon-Sicairos, A. Canizalez-Roman, A.R. dos Santos, S.S. Hassan, et al., Exoproteome and secretome derived broad spectrum novel drug and vaccine candidates in Vibrio cholerae targeted by Piper betel derived compounds, PLoS One 8 (2013), e52773. .
[84] E.W. Hewitt, The MHC class I antigen presentation pathway: strategies for viral immune evasion, Immunology 110 (2003) 163–169.
[85] E.Y. Jones, L. Fugger, J.L. Strominger, C. Siebold, MHC class II proteins and disease: a structural perspective, Nat. Rev. Immunol. 6 (2006) 271.
[86] E.H. Chung, Vaccine allergies, Clin. Exp. Vaccine Res. 3 (2014) 50–57.
[87] Z. Liu, Z. Wang, Y.-H. Chen, Predefined spacers between epitopes on a recombinant epitope-peptide impacted epitope-specific antibody response, Immunol. Lett. 97 (2005) 41–45.
[88] M.J. Motamedi, J. Amani, S. Shahsavandi, A.H. Salmanian, In silico design of multimeric HN-F antigen as a highly immunogenic peptide vaccine against Newcastle disease virus, Int. J. Pept. Res. Ther. 20 (2014) 179–194.
[89] T.E. Creighton, Disulphide bonds and protein stability, BioEssays 8 (1988) 57–63.
[90] A.A. Dombkowski, K.Z. Sultana, D.B. Craig, Protein disulfide engineering, FEBS Lett. 588 (2014) 206–212.
[91] R.K. Pandey, T.K. Bhatt, V.K. Prajapati, Novel immunoinformatics approaches to design multi-epitope subunit vaccine for malaria by investigating anopheles salivary protein, Sci. Rep. 8 (2018) 1125.
[92] F.U. Haq, A. Abro, S. Raza, K.R. Liedl, S.S. Azam, Molecular dynamics simulation studies of novel $β$-lactamase inhibitor, J. Mol. Graph. Model. 74 (2017) 143–152.
[93] V.N. Maiorov, G.M. Crippen, Significance of Root-Mean-square Deviation in Comparing Three-dimensional Structures of Globular Proteins, 1994.
[94] S. Abbasi, S. Raza, S.S. Azam, K.R. Liedl, J.E. Fuchs, Interaction mechanisms of a melatonergic inhibitor in the melatonin synthesis pathway, J. Mol. Liq. 221 (2016) 507–517.
[95] S. Ahmad, S. Raza, R. Uddin, S.S. Azam, Binding mode analysis, dynamic simulation and binding free energy calculations of the MurF ligase from Acinetobacter baumannii, J. Mol. Graph. Model. 77 (2017) 72–85, https://doi.org/10.1016/j.jmgm.2017.07.024.
[96] M.Y. Lobanov, N.S. Bogatyreva, O.V. Galzitskaya, Radius of gyration as an indicator of protein structure compactness, Mol. Biol. 42 (2008) 623–628.
[97] A.R. Fersht, The hydrogen bond in molecular recognition, Trends Biochem. Sci. 12 (1987) 301–304.
[98] R.E. Hubbard, M. Kamran Haider, Hydrogen Bonds in Proteins: Role and Strength, E LS, 2001.
[99] D.N. Kouetcha, H. Ramézani, N. Cohaut, Ultrafast scalable parallel algorithm for the radial distribution function histogramming using MPI maps, J. Supercomput. 73 (2017) 1629–1653.
[100] J. Donohue, Radial distribution functions of some structures of the polypeptide chain, Proc. Natl. Acad. Sci. 40 (1954) 377–381, https://doi.org/10.1073/pnas.40.6.377.
[101] I. Jelesarov, A. Karshikoff, Defining the role of salt bridges in protein stability, Protein Struct. Stability, Interact, Springer 2009, pp. 227–260.
[102] Z.S. Hendsch, B. Tidor, Do salt bridges stabilize proteins? A continuum electrostatic analysis, Protein Sci. 3 (1994) 211–226.
[103] H.R. Bosshard, D.N. Marti, I. Jelesarov, Protein stabilization by salt bridges: concepts, experimental approaches and clarification of some misunderstandings, J. Mol. Recognit. 17 (2004) 1–16.
[104] G.I. Makhatadze, V.V. Loladze, D.N. Ermolenko, X. Chen, S.T. Thomas, Contribution of surface salt bridges to protein stability: guidelines for protein engineering, J. Mol. Biol. 327 (2003) 1135–1148.